JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1977, 10, 231-238 NUMBER 2 (SUMMER) 1977

PEER TUTORING AMONG ELEMENTARY STUDENTS: EDUCATIONAL BENEFITS TO THE TUTOR[1]

JOHN P. DINEEN,[2] HEWITT B. CLARK, AND TODD R. RISLEY

UNIVERSITY OF KANSAS AND JOHNNY CAKE CHILD STUDY CENTER

To determine whether tutoring might be academically beneficial to the tutor, this study investigated the acquisition of spelling words by three elementary students in a peer tutoring program. The experimental design allowed a simultaneous comparison of each child's gain in performance on comparable word lists on which the child tutored another child, was tutored by another child, or neither gave nor received tutoring. The children's spelling improved nearly an equivalent amount on those words on which they tutored another child as on the words on which they were tutored; no such change was noted on the words on which they neither gave nor received tutoring. These findings, that peer tutoring is profitable for the tutor as well as the tutee, provide a basis for recommending peer tutoring as one method of individualizing education.

DESCRIPTORS: academic behavior, individualized instruction, tutoring, group consequences, primary classroom, methodology, experimental design, peers, students

Educators have employed a number of approaches to individualizing the learning process, among which have been teaching machines (Skinner, 1961), paraprofessional aides (Hanley and Perelman, 1971), and personalized, self-paced instruction (Keller, 1968). Tutoring represents still another approach to individualized instruction. One student teaching another is not a recent development (Horst, 1931), and a variety of tutorial programs have been suggested (*e.g.*, Ellson, Barber, Engle, and Kampwerth, 1965; Fleming, 1969; Lippitt, 1969; Neidermeyer and Ellis, 1971) and reviewed (Gartner, Kohler, and Riessman, 1971).

Several studies have suggested that students who are tutored (tutees) benefit academically from the procedure. Frager and Stern (1970) and Johnson and Bailey (1974) found that kindergarten students tutored in language readiness or mathematics skills by elementary students made significant gains over control groups that received no tutoring. Similarly, in studies by Harris, Sherman, Henderson, and Harris (1972) and Harris and Sherman (1973), fourth- and fifth-grade students were asked to arrange themselves in groups of two or three with their peers to work together on spelling words or math problems, but were not required to follow a specific tutorial procedure. Nor had the teacher appointed a specific tutor. While engaging in tutoring, these students made larger gains on their spelling or worked at a higher rate and with greater accuracy on math problems than when working independently. The unstructured classmates (peer) tutoring procedures used in the Harris *et al.* (1972), and Harris and Sherman (1973), studies suggest that students can benefit from being taught by, or perhaps by teaching

[1]This is one of a series of studies by the Living Environments Group of the University of Kansas. This research was supported by the Johnny Cake Child Study Center Foundation and by a NICHD grant (HD03144) to the Bureau of Child Research and the Department of Human Development of the University of Kansas. The authors wish to thank Jerry Davis and John Macrae for their advice in the preparation of this manuscript, and Becky Clark for her helpfulness and patience in the classroom. Reprints may be obtained from the Living Environments Group, c/o Todd R. Risley, Department of Human Development, University of Kansas, Lawrence, Kansas 66045, or from Hewitt B. "Rusty" Clark, Children's Behavioral Services, 6171 West Charleston Boulevard, Las Vegas, Nevada 89109.

[2]Now at Department of Rehabilitation Medicine RJ30, University Hospital, University of Washington, Seattle, Washington 98195.

others but those studies did not analyze how much a child learned as tutee, and how much as tutor.

Although the potential benefits of tutoring for the peer tutor have not been experimentally analyzed, one study using older students to tutor younger, less experienced students (cross-age tutoring) has suggested that the process is beneficial for both tutor and tutee. Cloward (1967) used tenth- and eleventh-grade students as tutors for fourth- and fifth-grade students whose reading achievement levels were below grade level. The results indicated that students who had been randomly assigned to be tutors scored statistically higher in reading skills than students (control) who did not participate in tutorial sessions but did, like the tutors, attend their normal classes in reading and language development.

While cross-age tutorial procedures appear to be a useful method of individualizing education, their implementation may pose certain practical problems. The older tutors, for example, often must either leave their classes or spend after-school hours participating in tutoring. If a tutorial procedure involving peers could demonstrate the same benefits to tutors and tutees as the cross-age study suggests, then the convenience of peer tutorial procedures, which would involve no after-school hours or movement of children from class, suggests that this is a more practical method of individualizing education.

Although the Cloward study suggests that the tutor can profit from tutoring, no evidence is available to indicate the same to be true of peer tutoring. Thus, it would be difficult to justify a student's time to tutor his peer. Yet, if tutoring appears to be a reasonable alternative to the traditional classroom lecture and individual studies approach to instruction, then it is critical to determine if the time students spend in tutoring their peers can be justified by demonstrating that they learn as they teach. In response to that issue, the present study was designed to evaluate experimentally the benefits of peer tutoring for the tutor, using elementary students who tutored spelling words to their peers.

METHOD

Subjects and Setting

Three children, Jane, Norman, and Brady, ages 9 to 10 yr, were members of an ungraded, open-environment classroom containing 12 children of normal intelligence, but with a 2-yr reading and a one-year math achievement deficiency. Two teachers were present.

All phases of the study were carried out in the classroom with the other children present. Tutorial sessions were held twice a day at 8:30 a.m. and 1:00 p.m. During each session, a tutor and tutee sat at a small table approximately 6 m from the closest of their other classmates, who were involved in other classroom activities with one of the teachers. The other teacher periodically monitored the tutorial sessions while also assisting several children working on an individualized mathematics program.

Procedure

Before any training began, third-grade spelling words were randomly selected to comprise six lists of 15 words each.[3] Each word was printed separately on a 10 by 15 cm index card, along with a picture, line drawing, or cartoon illustrating the word.

Tutor training. The procedures to be used by the tutor and tutee were taught to each child individually through a combination of modelling and role-playing techniques. A 20- to 30-min session was needed for each of the three children, during which the teacher initially modelled appropriate tutor behaviors, and then assumed the role of tutee while a child acted as tutor. The teacher praised appropriate tutor behaviors, such as slowly spelling a word to a tutee or making a tactful verbal correction. After the child displayed appropriate tutor behaviors, the teacher and child switched roles to allow the child to practise as tutee. The same prompting, model-

[3]The spelling words were taken from a standard third-grade text by K. Day and P. Lightfoot, *Words and patterns, "A" spelling series, level "B"*. Chicago: Science Research Associates, 1970.

ling, and shaping procedures were used. The criteria for ending the tutor training sessions were met when the child used the tutor and tutee procedures (described later) consistently over several trials.

Word identification. Before pretesting, each child was taught to identify the words on 45 index cards (three of the word lists). Initially, the teacher held up a card, said the word, and praised the child for repeating it correctly. Gradually, over trials, the teacher discontinued saying the words until the child could say the word as the cards were presented. The training was continued until each child individually could verbally identify at least 41 out of the 45 (91%) words on the cards without prompting. This training and testing procedure took about 40 min per child.

Spelling pretest. Immediately after reaching criterion on word identification over the 45 words (three lists), the three children took a tape-recorded spelling test in which these 45 words were randomly presented. On the tape, the word was said, used in short phrase or sentence, then said again. Every 20 sec a new word was presented. No talking was allowed among the children, and on-task behavior such as writing neatly on the test paper and keeping on the correct line of the page was praised. Following the test, the papers were corrected and the children told the number of words they had spelled correctly. They were allowed to choose a special activity from a reinforcement menu posted in the classroom with the number of words correctly spelled determining the types and durations of the activities that could be selected. The activities included access to a variety of play materials, playing outside, or receiving pieces of candy.

Tutoring. During the 20-min tutorial sessions the tutor and tutee sat facing each other. On the table were two boxes large enough to accept the 10- by 15-cm index cards. One box was marked with a "plus" and the second with a "minus". The tutor held all 15 cards (one of the three lists) so the tutee could not read them. The tutor said the word on the top card and waited for the tutee to attempt to write it on a blank piece of paper and then spell it aloud. If the tutee's written and oral spellings were both correct, the tutor circled the word on the paper with a red pen, praised the tutee, and dropped the card into the plus box. Only if the word was spelled correctly on the first presentation was the card dropped in the plus box. If either the written or oral spelling was incorrect, the tutor drew a single red line through the word on the paper and said that the word was misspelled or described the error, *e.g.,* said that a particular letter was missing. The tutor then slowly spelled out the word while the tutee wrote it out. Again the tutee spelled the word out loud and presented the written version to the tutor. If both the written and oral spellings were correct, the tutor circled the word, praised the tutee, and placed the card in the minus box. Otherwise, the tutor again slowly spelled the word while the tutee wrote it out. The tutee spelled out the word and presented it to the tutor, who checked the word, praised the tutee if the word was spelled correctly, and placed the card in the minus box. After the tutor placed a card in either the plus or minus box, the word on the next card was spoken. Therefore, when a word was presented, the tutee could (1) correctly spell the word, and go on to the next word; (2) misspell the word, be corrected, respell the word correctly and go on to the next word; or (3) misspell the word, be corrected, misspell the word again, be corrected, spell or misspell it again and go on to the next word. If a tutee refused to cooperate in spelling out or writing the word, then the tutor placed the card directly in the minus box and went on to the next word.

After the tutor presented all 15 cards to the tutee, the tutee's used paper was replaced with a blank one, thus preventing the tutee from referring to previous spellings of the words, and began anew with the cards that had been placed in the minus box. These cards were again placed only in the plus box when they were spelled correctly the first time. Cards from the minus box were used in each successive round thereafter un-

til all the words were placed in the plus box. At that time, the tutor would remove the 15 cards from the plus box and begin again with them. At the end of each 20-min session the cards in the minus box and those held by the tutor were set aside so the next session could resume where the previous tutoring session had ended.

Following each tutorial session, the tutee took a tape-recorded spelling test, in the same format as the pretest but only over the 15-word list that had been tutored. The test was immediately graded and both the tutor and tutee were told the number of words spelled correctly, and both were allowed independently to choose a special activity from the reinforcement menu based on the number of words correctly spelled by the tutee during the test. No information was given about which words were incorrectly spelled on the taped test.

During 20-min tutorial sessions, the children were unobtrusively monitored by the teacher, located with other students in the classroom. An average of 14.7 times per session, on a predetermined random (VI 1.3-min schedule) the teacher would observe the tutor and tutee for 10-sec. Then the teacher immediately walked over to the children and praised them for on-task behaviors such as writing or verbally spelling the words clearly, or fined them for being off-task. Both tutor and tutee were fined for off-task behavior even if only one was not working appropriately. Each fine resulted in a proportionate loss of either a morning or afternoon snack or recess. Of these interventions, 91.5% involved praise and 8.5% involved a fine. These figures did not vary substantially across subject pairs or series.

Posttest. Twenty-four hours after the eighth tutorial session, a spelling test containing the 45 words from the three lists was given to all three children. Identical procedures as well as the same tape-recorded test was used in both the pre- and posttests.

Experimental design. The independent variable for each child was their being tutor, tutee, or having no training on a list of words. The dependent variable involved a comparison between the number of words spelled correctly on a pretest and posttest composed of three word lists.

The children were exposed to two Series in the study. For each child, Series I included word identification training on three word lists (45 words), a pretest on the three lists, eight tutorial sessions as a tutor on one word list, eight tutorial sessions as tutee on a second word list, and a posttest over the three lists. Each child received no tutorial training on a third word list (control list). After Series I was completed for all children, Series II was begun. It included word identification training on three new word lists (45 words), a pretest on the three lists, eight tutorial sessions as a tutor on one word list, eight tutorial sessions as a tutee on a second word list, no tutoring on a third list, and a posttest over the three lists. As Figure 1 illustrates, during Series I, Jane tutored Norman on list 1, Norman tutored Brady on list 2, and Brady tutored Jane on list 3. In Series II, the order of tutoring was

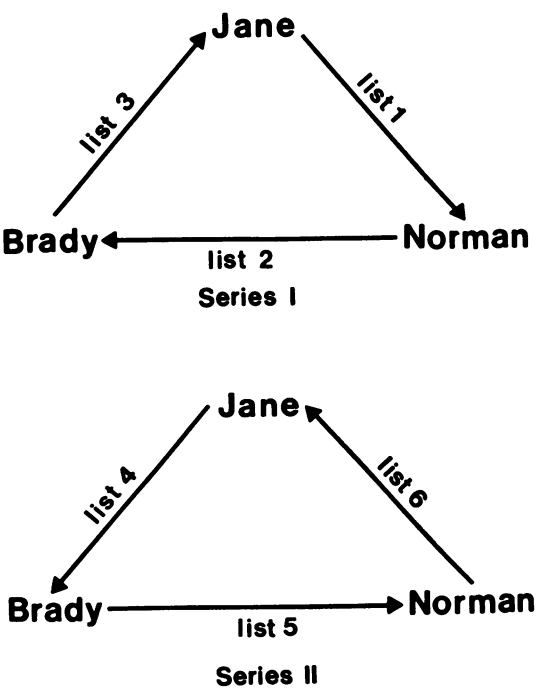

Fig. 1. The experimental design for this study. The arrows point from the tutor to the tutee; for example, in Series I Jane tutored Norman on word list 1. In Series II, however, Jane tutored Brady on word list 4.

reversed, and three new lists of words were used so that Jane tutored Brady on list 4, Brady tutored Norman on list 5, and Norman tutored Jane on list 6.

Thus, in each series, each of the three word lists served a different function for each child. That is, each word list served as the tutor word list for one child, as the tutee word list for another child, and as the control word list for a third child. Therefore, the term "tutor words" refers to those words a child taught, "tutee words" to those that were taught to him, and "control words" to those on which the child had no tutorial exposure.

Interobserver agreement. In both Series of pretests and posttests, each word was scored as correctly or incorrectly spelled by two observers working independently. For purposes of interobserver reliability, an agreement was defined as both observers scoring a word as correct, or both scoring the word incorrect; a disagreement was defined as only one observer scoring a word as correctly spelled.

Interobserver agreements for both Series were 536 agreements of 540 opportunities, or 99%.

RESULTS

The mean per cent change for all three children for both series was a loss of 1% on the control words, a gain of 59% on the tutee words, and a gain of 47% on the tutor words, as shown in Figure 2.

Figure 3 shows the results for tutor, tutee, and control words separately for each child and for both series of tutorial sessions. Jane, for example, began Series 1 with a pretest score of zero correct tutor words, two correct tutee words, and three correct control words. Each of the eight tutorial sessions in which she was a tutee was followed with a test over the tutee words. The results show a gradual acquisition in correct spelling of these words. On the posttest, Jane got nine tutor words correct for an improvement of 60% from the pretest, 13 tutee words correct for an improvement of 74%, and two control words correct for a decrement of 7% from pretest performance. The other children showed similar changes except that Norman did not learn as many words tutoring as he did being tutored.[4]

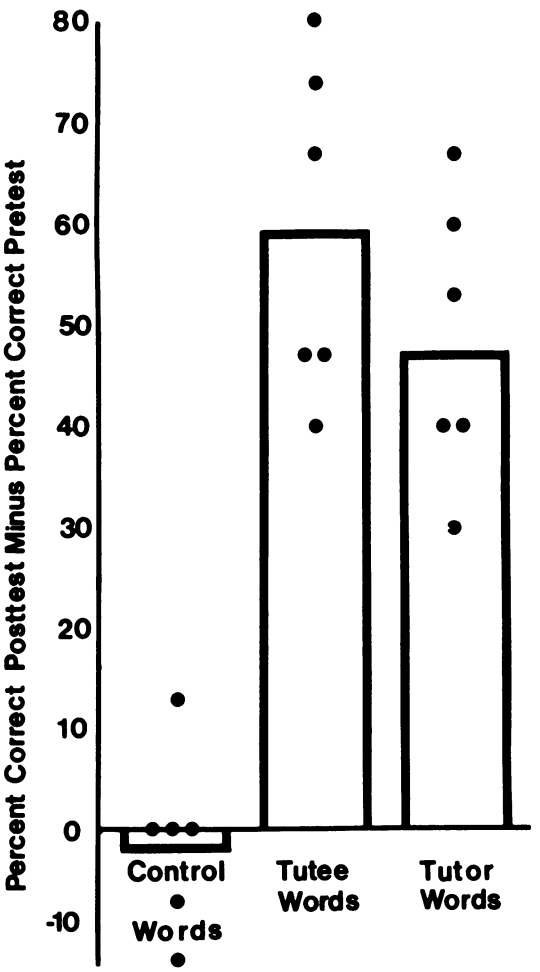

Fig. 2. Mean per cent changes in number of words spelled correctly from pretest to posttest for both Series for the control, tutee, and tutor words. The dots indicate the individual percentage changes for all three children for both Series.

[4]Since the variability surrounding each mean is directly displayed in Figures 2 and 3, and the differences are large and consistent, statistical analysis is not needed to reach a conclusion about these data. However, a consideration of the statistical analyses that can be applied to these data will exemplify the intra- and intersubject comparisons inherent in the "triadic" experimental design. Since the word lists were randomly assigned to children within each series and the order of tutoring counterbalanced between Series I and II, a t-test for correlated groups (McCall, 1970) could be

DISCUSSION

The results indicate that tutoring a peer increased the subjects' spelling accuracy nearly as much as being tutored by a peer; a no-practice condition produced little change in spelling accuracy, suggesting that the process of tutoring is not necessarily a waste of the tutor's time. Neither the effects of a maturation nor a test-retest effect in the absence of training would be likely to explain the tutor and tutee's gains, since such an effect was not observed on the list of control words. It appears, therefore, that the interaction between tutor and tutee during the tutorial process accounted for the substantial improvement in number of words spelled correctly by the tutor.

One possible factor in the tutor's marked improvement in spelling is that reinforcement for participating in a tutorial session was contingent on the tutee's performance in a spelling test after each tutorial session. The tutor might thus have been more careful to evaluate the tutee's spelling than if reinforcement was contingent on some other aspect of the tutorial session. Another important factor might be that of the structured tutorial procedure. The management system of frequent teacher intervention, and clearly specified behaviors that both children were to follow apparently encouraged both the tutor and tutee to attend closely to their tasks. The children learned to use the procedure relatively quickly,

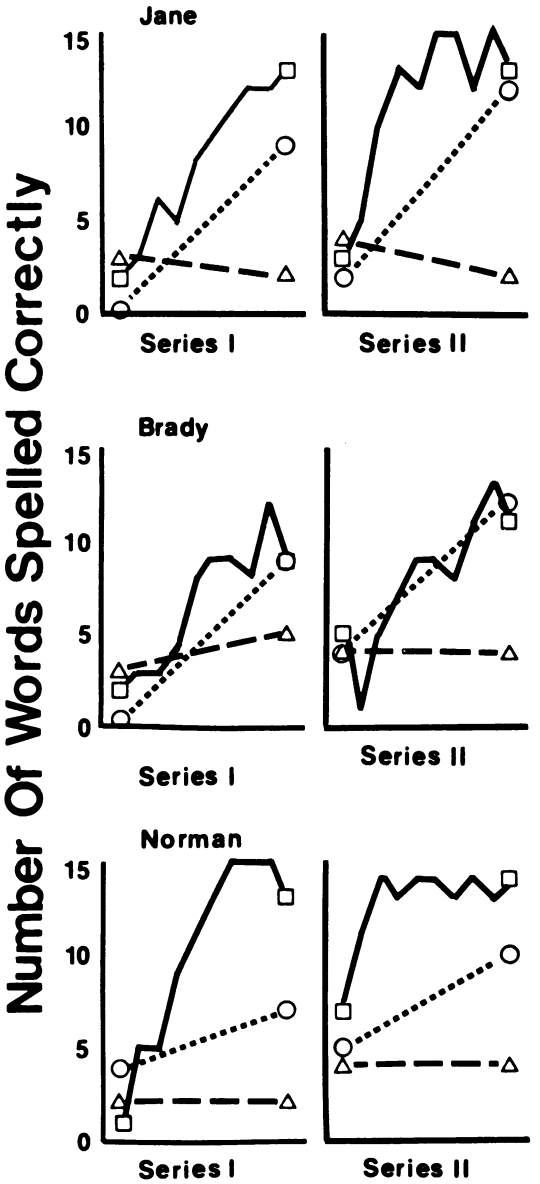

Fig. 3. Pretest, tutorial sessions tests, and posttest scores for all three children. Triangles connected by dashed lines indicate the pre- and posttest scores on words on which no training (control) was given. Squares connected by solid lines indicate the pretest, tutorial sessions tests, and posttest scores of tutee words. Circles connected by dotted lines indicate the pre- and posttest scores of tutor words.

used to analyze statistically the change in spelling performance from pre- to posttest relative to the *inter-subject* variability. For example, with the present data, such a comparison indicated significant differences ($p < 0.05$) between the tutee and control word performances ($df = 2$, $t = 3.65$), and no significant difference between the tutor and tutee performances ($df = 2$, $t = 1.52$). A t-test comparison could similarly be made of the average change scores relative to the total variability contributed by both subjects and series. However, since words were randomly assigned to word lists, and sessions on which each child was a tutor and tutee were interspersed, *intrasubject* statistical comparisons are possible. For example, with the present data, a simple Chi Square analysis (Underwood *et al.*, 1954) of the distribution of words learned by each subject between conditions in the two series combined indicated significant differences ($p < 0.005$) between tutor and control, and tutee and control conditions for each subject ($df = 1$, Chi Square > 8.00 in each case), and no significant differences between tutor and tutee conditions for Jane or Brady, but a significant difference ($p < 0.005$) for Norman.

and the teacher intervention data suggest that they maintained high levels of on-task behavior throughout the study. As part of the procedure, the tutor listened to the verbal and examined the written version of each spelling word, and frequently verbally spelled the word, therefore practising spelling the word nearly as often as the tutee. It is interesting to note, however, that the tutor never wrote a spelling word, nor took tests over the words he or she taught after each tutorial session.

The possible lack of uniformity among word lists was partially compensated for since any given list would have been distributed among all three children, with one child using it to tutor, a second child being tutored on it, and the third child using it as a control. Then too, the possibly differing academic abilities of children were minimized by having each child serve as a tutor, tutee, and control. Reversing the tutor-tutee roles from Series I to Series II further balanced the design allowing each child to tutor and to be tutored by each of the other children. The research design allowed for both inter- and intra-subject comparisons, since each child was his or her own control as well as a control for the other children. The advantages of this "triadic" design are that it conveniently allowed (1) a simultaneous comparison of the tutor, tutee, and no-practice functions; (2) a balancing of both subject and word list inequalities; and (3) a large number of comparisons over a short period of time with a relatively small number of subjects.

The use of student tutors appears to be an attractive alternative to large-group instruction, particularly if, as shown in the present study, it offers similar beneficial effects to the students giving and receiving tutoring. Participation in the tutorial sessions appeared to promote an increase in the subjects' cooperative social behaviors in other school settings and they expressed a preference for tutoring over independent study. However, no data were taken on these changes. For the teacher, the procedure allowed more time for helping students with academic problems, because monitoring the tutoring took less time than individuallly checking each child's work. Because students taught their peers, scheduling did not involve other teachers or school administrators, making implementation relatively easy. Although this study did employ a structured tutorial system that may be applicable for peer tutoring in elementary classrooms, the study did not examine all practical applications of the system. Yet, as educational practices become more student centered, and individualized instruction plays an increasingly important role in the teaching process, this demonstration provides some positive support for using peer tutors.

REFERENCES

Cloward, R. D. Studies in tutoring. *Journal of Experimental Education,* 1967, **36,** 14-25.

Ellson, D. G., Barber, L., Engle, T. L., and Kampwerth, L. Programmed tutoring: A teaching aid and research tool. *Reading Research Quarterly,* 1965, **1,** 77-127.

Frager, S. and Stern, C. Learning by teaching. *The Reading Teacher,* 1970, **23,** 403-417.

Fleming, J. C. Pupil tutors and tutees learn together. *Today's Education,* 1969, **58,** 22-24.

Gartner, A., Kohler, M. C., and Riessman, F. *Children teach children: learning by teaching.* New York: Harper & Row, 1971.

Hanley, E. M. and Perelman, P. F. Research resulting from a model cities program designed to train paraprofessionals to aid teachers in elementary school classrooms. In E. A. Ramp and B. L. Hopkins (Ed), *A new direction for education: behavior analysis,* 1971. Lawrence: The University of Kansas Support and Development Center for Follow Through, 1971. Pp. 158-190.

Harris, V. W. and Sherman, J. A. Effects of peer tutoring and consequences on the math performance of elementary classroom students. *Journal of Applied Behavior Analysis,* 1973, **6,** 587-597.

Harris, V. W., Sherman, J. A., Henderson, D. G., and Harris, M. Effects of peer tutoring on the spelling performance of elementary classroom students. In G. Semb (Ed), *Behavior analysis and education—1972.* Lawrence: The University of Kansas Support Development Center for Follow Through, 1972. Pp. 222-231.

Horst, H. M. History of student tutoring at West High School, Akron, Ohio. *Junior-Senior High School Clearinghouse,* 1931, **24,** 69-81.

Johnson, M. and Bailey, J. S. Cross-age tutoring: fifth graders as arithmetic tutors for kindergarten children. *Journal of Applied Behavior Analysis,* 1974, **7,** 223-232.

Keller, F. Good-bye teacher. *Journal of Applied Behavior Analysis,* 1968, **1,** 78-89.

Lippitt, P. Children can teach other children. *The Instructor,* 1969, **78,** 41, 99.

McCall, Robert B. *Fundamental statistics for psychology.* New York: Harcourt, Brace & World, Inc. 1970.

Neidermeyer, F. C. and Ellis, P. A. Remedial reading instruction by trained pupil tutors. *The Elementary School Journal,* 1971, April, 400-405.

Skinner, B. F. Why we need teaching machines. *Harvard Education Review,* 1961, **31,** 377-398.

Underwood, B. J., Duncan, C. P., Taylor, J. A., and Cotton, John W. *Elementary Statistics.* New York: Appleton-Century-Crofts, 1954.

Received 4 March 1795.
(Final acceptance 3 September 1976.)